What is Program Evaluation?

This paper attempts to explain what program evaluation is, what it's not, and to anticipate and answer some questions about it.

I.____Definitions

The General Accounting Office (GAO), which performs evaluations of various programs throughout the federal government, defines program evaluation as:

> "A systematic study conducted to assess how well a program is working...typically focused on achievement of program objectives."

The Office of Management and Budget (OMB) definition of program evaluation includes:

> "An assessment, through objective measurement and systematic analysis, of the manner and extent to which Federal programs achieve intended objectives."

Since these definitions focus in part on assessing success in meeting program goals and objectives, you may wonder how program evaluation differs from *performance measurement*, i.e., monitoring and reporting of program accomplishments. The answer is that performance measurement generally describes what a program achieved (e.g., outputs or outcomes) during a given time period, often in comparison to preestablished goals.

In contrast, program evaluation can help explain *why* these results occurred. Knowing *why* is the key to sustaining good performance or improving poor performance. Program evaluation can help managers and staff find out what's working well and what isn't. It can be used to identify areas needing improvement, strategies for better achieving established goals, and ways to improve data collection or measurement of program results. Performance measurement alone cannot answer these questions.

II.____Types of Program Evaluations

There are four principal types of program evaluations. The four types are not mutually exclusive, however. Elements of more than one type of evaluation may be found in a single evaluation project.

Outcome evaluation

Assesses the extent to which programs achieve their outcome-oriented objectives. Examines the reasons for differences between actual outputs/outcomes and stated objectives

(e.g., *why* the number and quality of permits issued exceeded or fell short of objectives). May include examination of program processes and activities to understand how outputs are produced or outcomes are achieved, and how quality and productivity could be improved.

Impact evaluation

A form of outcome evaluation that assesses the net effect of a program by comparing program outcomes with an estimate of what would have happened in the absence of the program. These evaluations help to isolate the program's contribution to achievement of an objective. Impact evaluations may prove especially useful when external, non-program factors are known to influence outcomes.

Process/implementation evaluation

Assesses the extent to which a program or process is operating as intended and identifies opportunities for streamlining or otherwise improving the program or process. Often begins with an analysis of how the process currently operates (e.g, What are the steps in permit issuance?). Also may assess the extent to which program operations/activities conform to statutory and regulatory requirements, agency policies, program design, relevant professional standards, or customer expectations.

Cost-benefit and cost-effectiveness analyses

Identifies program benefits, outputs, or outcomes and compares them with the internal and external costs (resources expended) to produce them.

III. Characteristics of Program Evaluation

In its broadest sense, program evaluation is simply a systematic way to learn from past experience. Program managers and staff learn from experience every day, of course. They constantly get anecdotal or statistical data that may help them form impressions about program effectiveness and efficiency. Such impressions are subject to many types of errors and biases and therefore are not necessarily a sound basis for decisions on future strategy, resource allocations, and so on. In contrast, a focused program evaluation study examines specifically identified factors of the existing program and goes broader and deeper than program managers and staff can plumb day-to-day.

Program evaluation is almost always retrospective, i.e., examining and learning from experience, though it may include prospective elements. For example, an analytical study that makes use of data on past performance to estimate future results would be an evaluation, but one done prospectively to estimate the effectiveness of a new environmental program based on assumptions about its design and/or operation would not be.

Program evaluation is always analytical (i.e., evaluators must make inferences, identify patterns, and reach conclusions from relevant data), but not all analytical studies are program evaluations. For example, a prospective analysis of the resources needed to implement a new environmental program would not be considered a program evaluation. Similarly, not all evaluations are program evaluations. For example, a technical evaluation of the performance of various contaminant clean-up techniques is not a program evaluation.

An evaluation can be systematic without being elaborate or expensive. It's possible to keep it simple and affordable while still systematically examining the program area in question. Doing it systematically does not necessarily mean asking every possible question and getting every bit of information anyone might want. Clearly identifying what you need to know and using logical ways of getting that knowledge make a study systematic.

A program evaluation need not cover the entirety of what EPA usually calls a program. Thus, it's possible to evaluate parts of, say, the RCRA program without evaluating the entire thing; indeed, it's usually more practical and instructive to focus on specific pieces rather than an entire program. Limiting the scope of the evaluation may enhance its chances for success.

Routine oversight or monitoring of Regional and state environmental programs (e.g., annual program reviews) are not by themselves program evaluations. The results of annual reviews of these programs could be used for evaluation, however, if the data are collected systematically and then consolidated to permit managers/staff to learn lessons that go beyond Region or state-specific performance issues. An in-depth evaluation of a Region's or state's implementation of a specific environmental program would likely be a program evaluation.

Program evaluations may include analysis of the adequacy of resources for an environmental program, but routine budget analyses or studies focused solely on resource needs are not program evaluations.

Projects aimed only at identifying and documenting 'success stories' and 'good practices' are not program evaluations. To have credibility, program evaluations must be objective and allow evaluators to examine all key issues. Program evaluations always include efforts to identify both what is working well and what isn't.

Studies undertaken solely to determine whether there has been fiscal waste, fraud, or abuse are not program evaluations.

IV. Who Conducts Program Evaluations?

A program evaluation may be conducted under the sponsorship or supervision of the program manager, i.e., by program staff or contractors, or it may be conducted independently by a third party. For example, some (but not all) studies done by GAO are program evaluations.

The National Academy of Public Administration (NAPA) has also participated in evaluation of EPA's programs.

Generally, if the results are going to be credible to the Congress or the public, an independent study is best. In most cases, however, program evaluations will be done primarily to provide insights and information that program managers and staff will use. This means that internal evaluations, often performed with contractor support, can be adequate to satisfy most program offices' needs.

For additional information on program evaluation, please visit:
http://intranet.epa.gov/evaluate/